8/1

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/020,895	12/19/2001	Soren Riis	1076.40981X00	8838

20457 7590 09/12/2005

ANTONELLI, TERRY, STOUT & KRAUS, LLP
1300 NORTH SEVENTEENTH STREET
SUITE 1800
ARLINGTON, VA 22209-3873

EXAMINER	
NG, EUNICE	
ART UNIT	**PAPER NUMBER**
2654	

DATE MAILED: 09/12/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/020,895	RIIS ET AL.	
	Examiner	Art Unit	
	Eunice Ng	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *19 December 2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-32* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-32* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *19 December 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *12/19/01,05/07//03*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Specification

1. The disclosure is objected to because of the following informalities: (1) page 11 of the specification refers to the "MLP phoneme counter 6" where element 6 is a mobile handset. Examiner has interpreted the MPL phoneme counter to be referring to element 23. (2) page 8 of the specification refers to the "mobile station 1" where element 1 is a speech pre-processor; the mobile station is not identified in the drawings.

Appropriate correction is required.

Information Disclosure Statement

2. The listing of references in the specification is not a proper information disclosure statement. 37 CFR 1.98(b) requires a list of all patents, publications, or other information submitted for consideration by the Office, and MPEP § 609 A(1) states, "the list may not be incorporated into the specification but must be submitted in a separate paper." Therefore, unless the references have been cited by the examiner on form PTO-892, they have not been considered.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4. Claims 1, 2, and 5-12 are rejected under 35 U.S.C. 102(b) as being anticipated by Gerber (paper on "A general approach to speech recognition," published Sept. 18-20, 1995).

Regarding claims 1 and 12, Gerber teaches a speech recognition system and method (see page 1, introduction section on speech recognition "machines" and methods of implementing the system), comprising means for and method of:

determining the length of a speech portion to be recognized (see page 5, section 4, lines 8-9, where "phoneme sequence" reads on a "speech portion");

defining a subset of speech portions for a set of stored speech portions in dependence on the determined length (see page 5, section 4, lines 9-11; see also Figure 3 on page 6 wherein "all words of length m±u" phonemes (a subset of speech portions) are taken from a "dictionary of size V phonemic transcribed words" (a set of stored speech portions)); and

recognizing the speech portion from the subset of speech portions (page 5, section 4, line 11-18, describes a sequence wherein a list of N-best matched words is outputted from the subset of speech portions, then a string (synonymous to a speech portion) matching algorithm is applied for recognizing the desired speech portion).

Regarding claim 10, Gerber teaches a speech recognition system, comprising:

a memory for storing a lexicon of speech portions (see page 3, section 3, line 1-2, discussing the use of the "TIMIT database" for experimental training; the speech database (reading on a "memory") was designed by DARPA to provide acoustic phonetic speech data for the development and evaluation of automatic speech recognition systems, and consists of utterances of 630 speakers that represent the major dialects of American English (reading on storing a lexicon of speech portions));

a counter arranged to determine the length of a speech portion to be recognized (see page 5, section 4, line 8, "endpoint detection algorithm");

a sub-lexicon definition module arranged to define a sub-lexicon from the lexicon of speech portions in dependence on the determined length (see page 5, section 4, lines 9-11, as specified in the rejection of claims 1 and 12; also Figure 3); and

a recognition module for recognizing the speech portion from the sub-lexicon of speech portions (see page 5, section 4, lines 11-18 as specified in the rejection of claims 1 and 12; also Figure 3).

Regarding claim 2, Gerber teaches wherein:

the subset defining means is arranged to define a subset of speech portions for each speech portion to be recognized (see page 5, section 4, lines 9-11; see also Figure 3 on page 6 wherein "all words of length m±u" phonemes (a subset of speech portions) are taken from a "dictionary of size V phonemic transcribed words" (a set of stored speech portions) for each speech portion to be recognized).

Regarding claim 5, Gerber teaches wherein:

the set of speech portions comprises a lexicon (synonymous to a dictionary) and the subset of speech portions comprises a sub-lexicon (synonymous to words extracted from the dictionary that are of length m±u phonemes). See Figure 3 and description on page 5, section 4, lines 9-11.

Regarding claim 6, Gerber teaches wherein:

the sub-lexicon comprises speech portions having a length similar to that of the speech portion to be recognized (page 5, section 4, lines 9-11 describe a sub-lexicon having "the length m of the recognized phoneme sequence ±u phonemes (similar in length) tolerance").

Regarding claim 7, Gerber teaches wherein:

the sub-lexicon comprises speech portions having a length which is the same as that of the speech portion to be recognized (as specified in claim 6, the sub-lexicon may have "±u phonemes tolerance," which reads on speech portions having a length which is the same, since "u" can be zero).

Regarding claim 8, Gerber teaches wherein:

the length of the speech portions in the sub-lexicon is determined in accordance with a confidence level associated with the length determining means (see page 9, second paragraph, particularly the last few lines which describe a confidence conditioned on a given interpretation (the phonemes) and time (which correspond to a valid sequence of phonemes, determined using the length determining means)).

Regarding claim 9, Gerber teaches wherein:

the speech portion comprises a word and the length determining means is arranged to detect the number of phonemes in the word (page 5, section 4, line 9).

Regarding claim 11, Gerber teaches a portable communications device comprising a speech recognition system (see third bullet under "Introduction," describing the use of the "SPHINX SR system," which is a speech recognizer developed under the Carnegie Mellon Sphinx Group that has been trialed in a range of applications (e.g. desktops, cell phones), which reads on implementation in portable communication devices).

5. Claims 13, 14, 17, 18, 21, and 25-29 are rejected under 35 U.S.C. 102(b) as being clearly anticipated by Russell et al. ("Measure of local speaking-rate for automatic speech recognition," published May 13, 1999).

Regarding claims 13 and 29, Russell et al. teach a speech recognition system in which an utterance to be recognized is represented as a sequence of phonetic segment models (see abstract, discussing "phone-level" speaking and estimation) in which a transition probability represents the probability of the occurrence of a transition between the models (see lines 4-5, "N-state HMM...

transition probability" under "ROS compensation"), comprising means (a speech recognizer) for, and a method of:

biasing the transition probabilities in dependence on the length of the utterance (see lines 9-10 under "ROS compensation," which discuss the state transition probabilities "scaled for fast speech," implying dependence on length).

Regarding claim 14, Russell et al. teach wherein the biasing means comprise means for applying a transition bias to each of the transition probabilities between a plurality of phonetic segment models (see lines 18-21 under "ROS compensation").

Regarding claim 17, Russell et al. teach means for estimating the number of phonetic segments in the utterance to be recognized (see lines 1-2 under "Phone-level measures of ROS" describing a measure of "phones-per-second" (or phonetic segments) in a sentence (synonymous with an utterance)).

Regarding claim 18, Russell et al. teach wherein the estimating means comprises a speaker specific rate of speech estimator (see Abstract).

Regarding claim 21, Russell et al. teach wherein the transition bias is set in response to the result of the estimating means (see lines 6-10 under "ROS compensation," which discuss a rate of speech compensation which scales (or biases) the state transition probabilities according to the speaker specific rate of speech).

Regarding claim 25, Russell et al. teach wherein the, or each, phonetic segment comprises a phoneme (see lines 1-2 under "Phone-level measures of ROS" describing "phone-level" measures wherein a "phone" is a sound unit of speech also known as phoneme, or allophone, which is predictable phonetic variant of a phoneme).

Regarding claim 26, Russell et al. teach a system wherein the, or each, utterance comprises a word (see line 3 under "Phone-level measures of ROS" describing phones "in a sentence," wherein a spoken sentence is a string of uttered words).

Regarding claim 27, Russell et al. teach wherein an utterance to be recognized is represented as a sequence of phonetic segment models in which a transition probability represents the probability of occurrence of a transition between the models (see lines 1-5, "N-state HMM... transition probability" under "ROS compensation"), comprising:

a phonetic segment estimator arranged to output an estimate of the number of phonetic segments in the utterance (see lines 1-2 under "Phone-level measures of ROS," wherein the utterance is a sentence); and

a processing module for applying a transition bias to the transition probability in response to the output of the estimator (see lines 6-10 under "ROS compensation," which discuss a rate of speech compensation which scales (or biases) the state transition probabilities according to the speaker specific rate of speech).

Regarding claim 28, Russell et al. teach a portable communications device including a speech recognition system (see line 16 under "experimental procedure," describing the use of a "DERA ASTREC speech recognizer," which is a state-of-the-art reconfigurable continuous automatic speech engine (or system) from The Defense Evaluation and Research Agency, which is suitable for deployment in command-and control direct voice input applications in a wide range of existing commercial markets (e.g. automotive, telephone-based IVR systems, TV control, etc.) and has already been trialed in a range of applications (e.g. European Fighter Aircraft), which reads on implementation in portable communication devices).

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

7. Claims 3 and 4 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gerber, as applied to claim 1, above, in view of Bergstrom et al. (US Patent No. 5,737,716).

Regarding claim 3, Gerber fails to teach a system wherein determining means comprises a neural network classifier. However, this feature is well known in the art as evidenced by Bergstrom et al., which disclose a neural network controlled speech analysis processor that includes a neural network which manages speech characterization, encoding, decoding, and reconstruction methodologies, reading on a neural network classifier (see abstract).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Gerber with those of Bergstrom et al. because Bergstrom et al. teach that this would "provide for rapid development, improved classification accuracy, improved speech analysis and speech synthesis architectures, and improved immunity to interference when trained with appropriate characteristic features" (see column 3, lines 15-19).

Regarding claim 4, Gerber also fails to teach a system wherein the neural network classifier comprises a multi-layer perceptron. However, this feature is well known in the art as evidenced by Bergstrom et al., which disclose a method and apparatus that implements "an advanced Multi-Layer Perceptron based structure" (see column 3, lines 4-6).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Gerber with those of Bergstrom et al., because Bergstrom et al. teach that this would "provide for improved speech synthesis, improved classification, improved robustness in interference conditions, improved bandwidth utilization, and greater flexibility" (see column 3, lines 7-11).

8. Claim 19 rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. in view of James et al. ("A Fast Lattice-Based Approach to Vocabulary Independent Wordspotting," ICASSP 1994, pp. 377-380).

Regarding claim 19, Russell et al. fail to teach a system wherein the estimating means comprises a Free Order Viterbi decoder. However, Viterbi decoders are well known in the field of speech recognition as evidenced by James et al., which disclose implementing a Free-Order Viterbi decoder (a null-grammar phone network, see page I-379, lines 14-15 of section 3.3).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Russell et al. with those of James et al., because James et al. teach that this would increase flexibility by being able to search for any word and speed of retrieval (see page I-377, sixth paragraph, lines 1-5; see also US Patent 6,073,095 to Dharanipragada et al. which references this publication in the "Prior Art" section of column 1).

9. Claim 20 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al., as applied to claim 17, above, in view of Bergstrom et al..

Regarding claim 20, Russell et al. fail to teach a system wherein the estimating means comprises a neural network classifier. However, this feature is well known in the art as evidenced by Bergstrom et al., which disclose a neural network controlled speech analysis processor that includes a neural network

which manages speech characterization, encoding, decoding, and reconstruction methodologies, reading on a neural network classifier (see abstract).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Gerber with those of Bergstrom et al., because Bergstrom et al. teach that this would "provide for rapid development, improved classification accuracy, improved speech analysis and speech synthesis architectures, and improved immunity to interference when trained with appropriate characteristic features" (see column 3, lines 15-19).

10. Claims 15, 16, and 30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. as applied to claims 14 and 29, above, in view of Gupta et al. (US Patent No. 5,390,278).

Regarding claims 15 and 16, Russell et al. fail to teach a system operable to recognize utterance from a recognition vocabulary, wherein the transition bias is calculated as the transition bias which maximizes recognition performance on a validation data set which represents, or has the same vocabulary as, the recognition vocabulary.

However, this procedure would have been obvious to one of ordinary skill in the art at the time the invention was made given the invention by Gupta et al.. Gupta et al. teach transition probabilities calculated, with "the one resulting in the best score" stored (see column 17, line 48-49), suggesting choosing a transition bias which maximizes recognition performance, and a validation data set representing, or having the same vocabulary as, the recognition vocabulary (see column 12, lines 45-49 and column 14, lines 21-23).

Regarding claim 30, Russell et al. fail to teach comprising decoding the sequence of phonetic segment models after application of the transition bias.

However, this procedure would have been obvious to one of ordinary skill in the art at the time the invention was made given the invention by Gupta et al.. Gupta et al. suggest decoding the sequence

of phonetic segment models after applying a bias (see Abstract and column 18, first paragraph; decoding is done by the A* search method as illustrated in Fig. 12a., element 418). Motivation for the combination would be to save the unnecessary decoding before the application of the transition bias, wherein the transition bias improves recognition.

11. Claim 31 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. as applied to claims 14 and 29, above, in view of Gupta et al. (US Patent No. 6,138,095).

Regarding claim 31, Russell et al. fail to teach comprising decoding the sequence of phonetic segment models without the application of transition bias (as specified in the rejection of claim 14, Russell et al. teaches only a transition bias) and normalizing the resulting scores by a contribution proportional to the transition bias.

However, this procedure would have been obvious to one of ordinary skill in the art at the time the invention was made given the invention by Gupta et al.. See column 3, lines 9-24 and column 3, line 66 through column 4, line 2 of Gupta et al. which discloses normalizing rejection thresholds and likelihood ratios (similar to resulting scores) by the magnitude of a null hypothesis probability (similar to transition probabilities). Motivation for the combination would be to simplify processing, in the case where the transition biases are too large, too small, or not integral numbers.

12. Claim 32 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell and Gupta et al. (US Patent No. 6,138,095), as applied to claim 31 above, further in view of Ueyama et al. (US Patent Application Publication 2001/0056346 A1).

Regarding claim 32, Russell et al. fail to teach comprising calculating the transition bias in parallel with the decoding of the sequence of phonetic segment models.

However, this procedure is well known in the art as evidenced by Ueyama et al., which disclose computing the output probabilities (synonymous to a transition probability) of acoustic models in parallel to decoding of speech parameters (synonymous with a sequence of phonetic segment models). See paragraph [0095]. Motivation for the combination would be to save time.

13. Claims 22-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. as applied to claim 21, above, in view of Schwartz et al. (US Patent No. 5,621,859), and further in view of Gupta et al. (US Patent No. 6,138,095).

Regarding claims 22-24, Russell et al. fail to teach a system comprising table look-up means for setting the transition bias in accordance with the number of phonetic segments in the utterance, and direct setting means for setting the transition bias as proportional or equal to the number of phonetic segments in the utterance.

However, a system comprising "table look-up means for setting the transition bias" is well known in the art as evidenced by Schwartz et al., which disclose a lookup-table where transition probabilities are stored for each transition from each grammar state to each possible following word (see column 15, lines 15-18 and 27-29; see also Figure 8). Motivation for the combination would be to reduce the amount of computation done by the system by storing transition probabilities already calculated.

Both Russell and Schwartz et al. fail to teach setting the transition bias in accordance with, or proportional to, the number of phonetic segments in the utterance.

However, setting the transition bias in accordance with, or proportional to, the number of phonetic segments in the utterance would have been obvious to one of ordinary skill in the art given the invention by Gupta et al.. Gupta et al. disclose that rejecting performance of speech recognition can be improved if a different rejection threshold is selected for each utterance length (see column 3, lines 46-

48), which is a synonymous to the idea of setting different transition biases that is utterance-length dependent or proportionally dependent, which includes setting the bias equal to the length. Gupta et al. teach that this would improve recognition performance for different utterance lengths (see column 1, line 58, through column 2, line 3).

Conclusion

14. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. This refers to US Patent No. 6,801,891 B2 to Garner et al., which disclose a speech recognition system with a method of decoding one or more sequences of sub-word units from a dictionary into one or more representative words; the US Patent No. 5,745,649 to Lubensky, which disclose automated speech recognition using a plurality of different multilayer perceptron structures to model a plurality of distinct phoneme categories; the US Patent No. 6,505,153 B1 to Van Thong et al., which disclose a speech rate calculation unit; the US Patent No. 6,539,353 B1 to Jiang et al., which disclose confidence measures using sub-word-dependence weighting of sub-word confidence scores for robust speech recognition; and the US Patent No. 5,638,487 to Chigier, which disclose recognizing speech represented by a sequence of frames of acoustic events separated by boundaries with assigned boundary probabilities conducted by a neural network.

15. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Eunice Ng whose telephone number is 571-272-2854. The examiner can normally be reached on Monday through Friday, 8:30 a.m. - 5:00 p.m..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Eunice Ng
AU 2654

24 Aug 2005

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER